REMARKS

Reconsideration of the application is respectfully requested in view of the foregoing amendments and following remarks. Claims 1-54 and 61-65 are pending in the application. No claims have been allowed. Claims 1, 29, 31, 33, 36, 47, 50, 54, 61, and 64 are independent. Various of the claims have been amended for clarification, and the scope of such claims is not thereby narrowed.

Information Disclosure Statement

Applicants note with thanks the Examiner's consideration of various references submitted in Information Disclosure Statements. Applicants note that pending U.S. Patent Application 09/407,021 to Chen et al. was submitted in an IDS but not considered by the Examiner. Applicants point out that a patent application, even if not published, can be the basis of a rejection. Accordingly, Applicants request the pending Application be considered.

Applicants also note an IDS citing two references (Chen et al. and Winston) was submitted with the Application, but the Form 1449 has not been returned by the Examiner to verify the references have been considered. For convenience of the Examiner, Applicants will submit the two references and the Chen patent application in an IDS to be filed shortly.

Drawing Correction

A marked-up copy of FIG. 6 is enclosed herewith to correct a typographical error. As shown, "FORM" should read "FROM." No new matter is added thereby. Applicants will submit a formal drawing upon approval of the correction by the Examiner.

Restriction

Applicants confirm election of the Group I claims and have canceled claims 55-60 without prejudice.

Cited Art

U.S. Patent No. 6,132,969 to Stoughton et al. ("Stoughton") is entitled "Methods for Testing Biological Network Models."

U.S. Patent No. 5,526,281 to Chapman et al. ("Chapman") is entitled "Machine-Learning Approach to Modeling Biological Activity for Molecular Design and to Modeling other Characteristics."

U.S. Patent No. 5,769,074 to Barnhill et al. ("Barnhill") is entitled "Computer Assisted Methods for Diagnosing Diseases."

Patentability of Claims 1-54 and 61-63 over Stoughton, Chapman, and Barnhill under § 103

The Action rejects claims 1-54 and 61 under 35 U.S.C. § 103(a) as unpatentable over Stoughton, Chapman, and Barnhill. Applicants respectfully submit the claims in their present form are allowable over the cited art. To establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. (MPEP § 2142.) As amended, the claims have one or more limitations not taught or suggested by the prior art.

Claim 1

As amended, claim 1 is directed to a method for quantifying gene relatedness. Applicants have amended the claim to recite "presenting a plurality of the quantifications of gene relatedness showing relative relatedness for a plurality of the permutations of the genes." For example, the Application shows example presentations at FIGS. 11 and 25. Further, the Application explains at page 12, line 27 - page 13, line 2:

> For the model 704, a value of 0.74 is provided and quantifies the effectiveness of the model 704 and the relatedness of the genes G_1, G_2, and G_3 (and the condition C_1). In the example, a high value indicates more relatedness than a low value, and the value falls between 0 and 1. However, any number of other conventions can be used (e.g., a percentage or some other rating).

The Action rejects claim 1, citing Stoughton. As understood by Applicants, Stoughton does not describe any arrangement teaching or suggesting the "presenting a plurality of the quantifications . . . showing relative relatedness" language of claim 1. Therefore, Stoughton does not teach or suggest the arrangement of claim 1.

The passages of Stoughton describing goodness of fit and a network model do not teach or suggest "showing relative relatedness" as recited in claim 1. For example, the Action relies on the following passage of Stoughton at column 10, lines 29-30:

> Preferably, the quantitative comparison returns a numerical value indicative of the goodness of the overall fit.

Stoughton goes on at column 10, lines 31-40 to say:

> In a preferred embodiment, quantitative comparison 105 further includes statistical tests of the significance of the goodness of the overall fit found for the network model. See, infra., Section 5.3. These tests construct an empirical probability distribution against which to test the actual network model against a null hypothesis, which is that the network model has no relation to biological system 104. These tests yield a numerical probability value (also called herein a "P-value") that the network model has no relation to the biological systems.

Thus, Stoughton does describe "goodness of fit," and "a numerical probability value . . . that the network model has no relation to biological systems." However, discussion of whether a network model has a relation to biological systems would not lead one to the claimed arrangement of "showing relative relatedness" as recited in claim 1.

Therefore, the passage in Stoughton does not teach or suggest the recited language because one of skill in the art could not be expected to surmise "showing relative relatedness" from a description of goodness of fit or whether a network model has a relation to biological systems.

Chapman fails to provide sufficient additional description for teaching or suggesting the claimed arrangement.

The Office also relies on Chapman, which describes in the Abstract:

> Explicit representation of molecular shape of molecules is combined with neural network learning methods to provide models with high predictive ability that generalize to different chemical classes . . .

Thus, Chapman does describe "neural network learning methods." However, such a description, alone, or in combination with Stoughton would not lead one of skill in the art to the recited "showing relative relatedness."

Barnhill also fails to provide additional description sufficient to teach or suggest the claimed arrangement.

The Office also relies on Barnhill, which describes at column 26, lines 9-12:

> 1. If the total number of diagnostic groups equals 2, go to next step. Otherwise, based on known facts about the disease process, organize the separation of groups into a binary classification tree.

Thus, Barnhill does describe "diagnostic groups." However, such a description, alone, or in combination with Stoughton and Chapman would not lead one of skill in the art to the recited "showing relative relatedness."

For these reasons, Applicants respectfully request the § 103 rejection be withdrawn. Applicants thus believe claim 1 and its dependent claims, 2-28 and 62-63, are allowable at this time.

Dependent claims 2-28 and 62-63

Without belaboring the individual language of the independent claims, Applicants note that the dependent claims recite various novel and nonobvious combinations not taught or suggested by Stoughton, Chapman, and Barnhill.

Claims 29-30

Independent claim 29 recites in part:

> estimating a coefficient of determination for sets of predictive elements and the target gene by comparing results of the multivariate nonlinear predictors with gene expression level observations for the target gene, wherein the predictive elements comprise expression level observations for genes other than the target gene; . . .

For example, the Application describes at page 3, lines 26-27:

> Effectiveness of the nonlinear model can be measured by estimating a coefficient of determination.

Claim 29 stands rejected over Stoughton. Applicants respectfully disagree. Applicants have obtained an electronic copy of Stoughton and performed a text search for "coefficient of determination" and found no mention of the language in Stoughton.

Stoughton does describe at column 29, lines 18-25:

> Further, to represent the two criteria on states in order to distinguish output classes, let coefficients $SD_{ij\alpha}$ be 1 if output classes i and j have different behaviors in state α, and be 0 otherwise; and let coefficients $SS_{ij\alpha}$ be 1 if output classes i and j have the same behaviors in state α, and be 0 otherwise. Both coefficients SS and SD can be generated by routine examination of the input/output table.

However, as understood by Applicants, the coefficients SS and SD are not and would not lead one of ordinary skill in the art to the recited language of "coefficient of determination."

As understood by Applicants, Chapman and Barnhill fail to include sufficient description that, in combination with Stoughton, would teach or suggest the recited "coefficient of determination."

For these reasons, claim 29 and its dependent claim, 30, are allowable over the cited art.

Claims 31-32

Claims 31-32 include the language "coefficient of determination" are therefore allowable for at least the same reasons stated for claims 29-30 above.

Claims 33-35

Independent claim 33 is directed to a method for identifying related genes and recites "testing effectiveness of the artificial intelligence function . . . to rate relatedness of the predicted gene and at least one gene associated with the predictive elements." Claim 33 stands rejected over a Stoughton-Chapman-Barnhill combination. Applicant respectfully disagrees.

Applicants do not understand any of these three references to describe "to rate relatedness of the predicted gene" or provide any way to modify the described systems so they would perform such an act. Thus, the three patents, alone or combined, do not teach or suggest "to rate relatedness of the predicted gene."

For these reasons, claim 33 and its dependent claims, 34-35, are allowable over the three patents.

Claims 36-46

Independent claim 36 is directed to a method of presenting an analysis of the expression levels to assist in identifying related genes and recites in part:

> for the predicted gene, constructing a plurality of nonlinear multivariate models predicting expression of the observed gene, wherein the nonlinear multivariate models comprise a variety of predictive elements chosen from permutations of expression levels of observed genes other than the predicted gene.

Claim 36 stands rejected over Stoughton, but Applicants respectfully disagree. Stoughton does not describe "wherein the nonlinear multivariate models comprise a variety of predictive

elements chosen from permutations of expression levels" as recited in claim 36. Stoughton does describe at column 10, lines 29-40:

> Preferably, the quantitative comparison returns a numerical value indicative of the goodness of the overall fit.
>
> In a preferred embodiment, quantitative comparison 105 further includes statistical tests of the significance of the goodness of the overall fit found for the network model. See, infra., Section 5.3. These tests construct an empirical probability distribution against which to test the actual network model against a null hypothesis, which is that the network model has no relation to biological system 104. These tests yield a numerical probability value (also called herein a "P-value") that the network model has no relation to the biological systems.

Thus, Stoughton does describe "a numerical probability value . . . that the network model has no relation to biological systems." However, such a description would not lead one to the claimed arrangement of "a variety of predictive elements chosen from permutations of expression levels" as recited in claim 36. Therefore, the description in Stoughton does not teach or suggest the language recited by claim 36.

Further, as understood by applicants, Chapman and Barnhill do not describe such an arrangement or how Stoughton could be modified to result in such an arrangement.

For these reasons, claim 36 and its dependent claims, 37-46, are allowable over Stoughton, Chapman, and Barnhill.

Claims 47-49

Independent claim 47 recites "wherein the nonlinear multivariate models have a variety of predictive elements chosen from permutations of expression levels." For at least the same reasons given for claims 36-46, above, the language of claim 47 is not taught or suggested by Stoughton or a Stoughton-Chapman-Barnhill combination.

For these reasons, claim 47 and its dependent claims, 48-49, are allowable over Stoughton, Chapman, and Barnhill.

Claims 50-53

Applicants have amended claim 50 to include "nonbinary." Claim 50 recites in part:

> means for constructing a nonbinary, nonlinear model predicting gene expression based on data comprising the plurality of gene expression level observations for the plurality of candidate genes;

Stoughton does describe at column 3, lines 12-16:

> One important abstraction involves coarsening the inputs and outputs of a network model to binary values and modeling interactions among cellular constituents in a biological system as logical gates with simple rules of combination.

However, as understood by Applicants, Stoughton fails to teach or suggest "constructing a nonbinary, nonlinear model predicting gene expression" as recited by claim 50.

Further, Chapman describes at column 4, lines 61-66:

> A novel modeling approach is proposed using a surface-based representation of molecular shape that employs neural network learning techniques to derive robust predictive models. Trained models predict the bioactive shape of molecules and can be readily interpreted to guide the design of new active compounds.

Thus, Chapman does describe "models predict the bioactive shape of molecules." However, one of skill in the art, presented with a description of predicting the bioactive shape of a molecule would not be led to the claimed arrangement involving "predicting gene expression." Further, as understood by Applicants, neither Chapman nor Stoughton contains any motivation for somehow combining the references to result in such an arrangement.

Finally, as understood by Applicants, Barnhill does not contain sufficient additional disclosure so that a Stoughton-Chapman-Barnhill combination would result in the claimed arrangement.

Amended claim 50 and its dependent claims, 51-53, are therefore allowable.

Claim 54

Claim 54 has been amended to clarify "to present the quantification as indicating relatedness for the plurality of candidate genes." The application shows an example at FIG. 11. As understood by Applicants, none of the references describes "to present the quantification as indicating relatedness" as recited in claim 54; further there is no motivation in the references to somehow modify or combine them to result in the claimed arrangement.

For these reasons, claim 54 is allowable over Stoughton, Chapman, and Barnhill.

Claims 55-60

Claims 55-60 have been canceled without prejudice.

Claim 61

Claim 61 recites "displaying a ranked list of gene relatedness." As understood by Applicants, a Stoughton-Chapman-Barnhill combination fails to teach or suggest such an arrangement.

Rejection of Claims 1-30, 33-54, and 61 § 112, Second Paragraph

Claims 1-30, 33-54, and 61 stand rejected under § 112, second paragraph, as indefinite. Specifically, the Action points to the words "related" or "relatedness" and assert the words can have two possible different meanings: evolutionarily linked genes or genes that code for proteins that interact in a particular pathway. Applicants respectfully disagree that the claims are indefinite.

First, Applicants refer to FIG. 4 of the Application, which is described at page 11, lines 18-24 as follows (emphasis added):

> The system 402 responsible for expression of a particular gene can be represented as shown in FIG. 4. In the system 402, a set of observable inputs X_1-X_m and a set of unknown inputs U_1-U_n operate to produce an observed result, $Y_{observed}$ (or simply Y). Some of the observable inputs X_1-X_m can correspond to gene expression levels. *For the sake of example, it is assumed that the internal workings of the system 402 are beyond current understanding.*

Also, the Application states at page 6, lines 21-23, "The mechanism of the relationship need not be a factor in determining relatedness." Thus, Applicants point out that relatedness can be quantified (and such quantification can be useful), even if the nature of the relationship is unknown. Hence it is unimportant whether the relationship is evolutionarily linked or linked by expression. For example, the workings of the system causing relatedness could be beyond current understanding and still exhibit relatedness. Identifying the related genes can be used, for example, to identify topics for further research to determine how the genes are related.

Second, Applicants refer the Examiner to the following examples of gene relatedness in the Application at page 6, lines 20-26:

> *Gene relatedness* includes genes having any of a variety of relationships, including coexpressed genes, coregulated genes, and codetermined genes. The mechanism of the relationship need not be a factor in determining relatedness. In the network that controls gene expression, a gene may be upstream or downstream from others; some may be upstream while others are downstream; or they may be distributed

> about the network in such a way that their relationship is based on chains of interaction among various intermediate genes or other mechanisms.

Accordingly, there is sufficient disclosure in the Application to enable one of skill in the art to determine the meaning of "relatedness."

The language "related" or "relatedness" is therefore sufficiently definite to particularly point out and distinctly claim the subject matter for purposes of § 112, second paragraph.

Patentability of New Claims 64-65

Claims 64-65 recite combinations not taught or suggested by the cited art. For example, claim 64 recites, "presenting at least one of the permutations of genes as related."

For these reasons, claims 64-65 are allowable.

Request For Interview

The Examiner is formally requested to contact the undersigned attorney prior to issuance of the next Office Action in order to arrange a telephonic interview. It is believed that a brief discussion of the merits of the present application may expedite prosecution. Applicants submit the foregoing formal Amendment so that the Examiner may fully evaluate Applicants' position, thereby enabling the interview to be more focused.

This request is being submitted under MPEP § 713.01, which indicates that an interview may be arranged in advance by a written request.

Conclusion

The claims in their present form should now be allowable. Such action is respectfully requested.

Respectfully submitted,

KLARQUIST SPARKMAN CAMPBELL
LEIGH & WHINSTON, LLP

By G. L. Maurer
Gregory L. Maurer
Registration No. 43,781

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

Marked-up Version of Amended Claims
Pursuant to 37 C.F.R. §§ 1.121(b)-(c)

1. **(Amended)** A computer-implemented method for quantifying gene relatedness for a plurality of candidate genes for which a plurality of gene expression level observations have been collected, the method comprising:

for a plurality of selected permutations of the plurality of candidate genes, performing (a)-(c):

(a) based on data comprising the plurality of gene expression level observations for the plurality of candidate genes, constructing a nonlinear model predicting gene expression **for the permutation of the plurality of candidate genes**;

(b) predicting gene expression with the nonlinear model; and

(c) measuring effectiveness of the nonlinear model in predicting gene expression, ~~**thereby quantifying**~~ **the effectiveness being a quantification of** gene relatedness for the plurality of candidate genes **of the permutation; and**

presenting a plurality of the quantifications of gene relatedness showing relative relatedness for a plurality of the permutations of the genes.

2. (Unchanged) A computer-readable medium comprising computer-readable instructions for performing the method of claim 1.

3. The method of claim 1 wherein the nonlinear model accepts a plurality of predictive elements as inputs, wherein at least one of the predictive elements indicates whether a gene expression observation is associated with having applied a particular external stimulus to biological material.

4. The method of claim 1 wherein the nonlinear model accepts a plurality of predictive elements as inputs, wherein at least one of the predictive elements indicates whether a gene expression observation is associated with a particular cell state.

5. The method of claim 1 wherein the nonlinear model accepts a plurality of predictive elements as inputs, wherein at least one of the predictive elements indicates differential gene expression between two samples of biological material.

6. The method of claim 1 wherein the nonlinear model comprises a multivariate prediction function accepting two or more inputs to predict gene expression.

7. The method of claim 1 wherein measuring effectiveness of the nonlinear model comprises comparing observed gene expression to gene expression predicted by the nonlinear model.

8. The method of claim 1 wherein constructing the nonlinear model predicting gene expression comprises choosing a nonlinear model from a constrained set of nonlinear models.

9. The method of claim 1 wherein measuring the model's effectiveness comprises evaluating the model to estimate a coefficient of determination for an optimal model estimated by the model.

10. The method of claim 1 wherein the nonlinear model predicting gene expression is a full-logic model predicting gene expression for a predicted candidate gene, and the effectiveness of the model is measured by comparing predictions of gene expression for the predicted candidate gene by the model with observations of gene expression for the predicted candidate gene.

11. The method of claim 1 further comprising:

obtaining the data comprising the plurality of gene expression level observations from results of a plurality of cDNA microarray experiments measuring mRNA transcription levels for a plurality of genes in biological material.

12. The method of claim 1 wherein the data comprising a plurality of gene expression observations is divided into a training set of data and a test set of data, wherein

the nonlinear model predicting gene expression is generated via the training set data; and

effectiveness of the nonlinear model is measured via the test set of data.

13. The method of claim 12 wherein the training set of data is extended by randomly reordering and recycling gene expression observations.

14. The method of claim 1 wherein

a plurality of training data sets are repeatedly chosen from the data comprising a plurality of gene expression observations;

the nonlinear model is one of a plurality of models constructed from the plurality of training data sets; and

quantification of the relatedness for the plurality of candidate genes is measured by measuring average effectiveness of the plurality of models constructed from the plurality of training data sets.

15. The method of claim 1 further comprising:

to determine contribution of a predictive element to the quantification of relatedness, constructing an additional nonlinear model predicting gene expression, wherein the additional nonlinear model has a single input.

16. The method of claim 1 wherein the nonlinear model predicting gene expression is a truth table predicting a gene expression level for a predicted candidate gene from predictive elements comprising expression level observations for candidate genes other than the predicted candidate gene.

17. The method of claim 16 wherein the truth table comprises ternary discrete values.

18. The method of claim 16 wherein gene expression levels in the truth table are ternary discrete values.

19. The method of claim 16 wherein

the truth table comprises a plurality of rows for possible combinations of expression level observations for the candidate genes other than the predicted candidate gene; and

for at least one of the rows, the truth table indicates predicted gene expression for the predicted candidate gene with a thresholded weighted average of gene expression level observations associated with the row.

20. The method of claim 1 wherein the nonlinear model predicting gene expression is a neural network predicting gene expression.

21. The method of claim 20 wherein the neural network consists of one neuron which predicts a gene expression level for a single predicted candidate gene.

22. The method of claim 21 wherein the neuron is a ternary perceptron accepting predictive elements as inputs, wherein the predictive elements comprise gene expression levels indicated as one of three possible values: up, unchanged, and down.

23. The method of claim 22 further comprising:

displaying a three-dimensional graph representing the ternary perceptron with two planes separating points on the graph into points relating to like predicted values.

24. The method of claim 22 further comprising:

displaying a three-dimensional graph representing the ternary perceptron with objects at points in three-dimensional space within the graph, wherein axes of the graph relate to thresholded gene expression levels for three of the candidate genes.

25. The method of claim 24 wherein the objects are of a color indicating a predicted gene expression level.

26. The method of claim 24 wherein the objects are of a size indicating a number of observations related to a point on the graph.

27. The method of claim 1 wherein the data comprising a plurality of gene expression observations comprises gene expression level observations generated by subjecting sample biological material to an experimental condition and observing regulation of mRNA transcription levels for a plurality of genes in the biological material as a result of being subjected to the experimental condition.

28. The method of claim 27 wherein the data comprising a plurality of gene expression observations further comprises an indication of the experimental condition to which the biological material related to an observation was subjected and the indication is included in the model to predict gene expression.

29. A computer-implemented method for identifying genes related to a target gene by analyzing gene expression level observations for the genes, the method comprising:

based on the gene expression level observations, constructing multivariate nonlinear predictors that predict an expression level for the target gene, wherein the predictors accept gene expression levels for other genes as predictive elements;

estimating a coefficient of determination for sets of predictive elements and the target gene by comparing results of the multivariate nonlinear predictors with gene expression level observations for the target gene, wherein the predictive elements comprise expression level observations for genes other than the target gene; and

ranking the groups of genes other than the target gene by coefficient of determination to present the genes other than the target gene in order of likelihood of relatedness to the target gene.

30. The method of claim 29 further comprising:

indicating a proper subset of the genes having the highest likelihood of relatedness to the target gene.

31. A computer-implemented method for analyzing gene expression level observations for a set of genes comprising a target gene, the method comprising:

estimating a coefficient of determination for an optimal multivariate nonlinear model predicting gene expression of the target gene by constructing a multivariate nonlinear model from the gene expression level observations of gene expression for the target gene, wherein the optimal multivariate nonlinear model and the constructed multivariate nonlinear model predict gene expression of the target gene based on variables representing gene expression levels of genes other than the target gene.

32. The method of claim 31 wherein the optimal multivariate nonlinear model and the constructed multivariate nonlinear model predict gene expression based, at least in part, on inputs comprising an indication of a condition to which biological material relating to the observations has been subjected.

33. A method for identifying related genes out of a set of genes for which gene expression level observations have been collected, the method comprising:

for at least one predicted gene out of the set of genes, training an artificial intelligence function to predict gene expression for the predicted gene, wherein the artificial intelligence function takes one or more predictive elements as inputs and produces a gene expression level for the predicted gene as an output, wherein at least one of the predictive elements is a gene expression level for a gene other than the predicted gene; and

testing effectiveness of the artificial intelligence function in predicting expression of the predicted gene to rate relatedness of the predicted gene and at least one gene associated with the predictive elements.

34. The method of claim 33 wherein the artificial intelligence function takes a plurality of predictive elements as inputs.

35. The method of claim 33 wherein the predictive elements comprise a variable indicating biological material was subjected to an experimental condition.

36. For a plurality of observed genes for which expression levels have been observed, a method of presenting an analysis of the expression levels to assist in identifying related genes, the method comprising:

denoting a particular observed gene as a predicted gene;

for the predicted gene, constructing a plurality of nonlinear multivariate models predicting expression of the observed gene, wherein the nonlinear multivariate models comprise a variety of predictive elements chosen from permutations of expression levels of observed genes other than the predicted gene;

measuring effectiveness of the nonlinear multivariate models in predicting expression of the predicted gene to quantify relatedness between the predicted gene and the set of genes associated with the predictive elements of the models; and

presenting a quantification of relatedness between the predicted gene and a set of genes associated with the predictive elements of at least one of the models.

37. The method of claim 36 further comprising:

for a set of predictive elements and a predicted gene, displaying a graph indicating the amount of increase in the effectiveness of the model for each of the predictive elements.

38. The method of claim 36 wherein at least two of the plurality of nonlinear multivariate models predicting expression of the observed gene are implemented in specialized hardware circuits for predicting gene expression.

39. The method of claim 36 further comprising:

displaying a user interface for evaluating the analysis, wherein the user interface comprises display elements graphically indicating the relatedness of the predicted gene to a plurality of gene sets.

40. The method of claim 39 further comprising:

displaying only those display elements indicating sets of genes in which each gene in the set improves the relatedness.

41. The method of claim 39 further comprising:

accepting as input a set of one or more designated predictor genes;

accepting as input a threshold relatedness;

accepting as input a set of one or more designated predicted genes; and

limiting the display elements of the user interface to those sets of genes having as members the one or more designated predictor genes and having at least the threshold relatedness for the one or more designated predicted genes.

42. The method of claim 39 further comprising:

accepting as input a set of one or more designated predictor genes; and

limiting display elements of the user interface to those sets of genes having the one or more designated predictor genes.

43. The method of claim 42 further comprising:

accepting as input a threshold increase in relatedness; and

further limiting display elements of the user interface to those sets of genes for which addition of the one or more designated predictor genes increases the relatedness by at least the threshold increase in relatedness.

44. The method of claim 36 further comprising presenting a ranking of gene sets according to their relatedness, wherein the ranking indicates which genes are in the sets.

45. The method of claim 44 wherein the ranking further indicates contribution of individual predictive elements to the effectiveness of the models.

46. The method of claim 36 wherein the predictive elements comprise a variable indicative of whether biological material related to a gene expression observation has been subjected to a particular condition.

47. For a plurality of observed genes for which expression levels have been observed, a method of performing an analysis of the expression levels to assist in identifying related genes, the method comprising:

(a) for a plurality of the observed genes, denoting a particular observed gene as a predicted gene and performing at least (b) and (c);

(b) for the predicted gene, constructing a plurality of nonlinear multivariate models predicting expression of the predicted gene, wherein the nonlinear multivariate models have a variety of predictive elements chosen from permutations of expression levels of observed genes other than the predicted gene;

(c) measuring effectiveness of the nonlinear multivariate models in predicting expression of the predicted gene to provide a quantification of relatedness between the predicted gene and genes associated with the predictive elements of the models.

48. The method of claim 47 further comprising:

skipping designating genes having fewer than a defined number of changes in expression level as predicted genes.

49. The method of claim 47 further comprising:

displaying a user interface comprising display elements indicating gene relatedness for a plurality of genes associated with the predictive elements for a plurality of predicted genes.

50. **(Amended)** A system for quantifying gene relatedness for a plurality of candidate genes for which a plurality of gene expression level observations have been collected, the system comprising:

means for constructing a **nonbinary,** nonlinear model predicting gene expression based on data comprising the plurality of gene expression level observations for the plurality of candidate genes;

means for predicting gene expression with the **nonbinary,** nonlinear model; and

means for measuring effectiveness of the **nonbinary,** nonlinear model in predicting gene expression, ~~thereby quantifying~~ **the effectiveness indicating** gene relatedness for the plurality of candidate genes.

51. The method of claim 50 wherein the means for predicting gene expression is a specialized hardware circuit.

52. The method of claim 50 wherein the means for predicting gene expression is a decision tree.

53. The method of claim 50 wherein the means for predicting gene expression is a truth table chosen from a constrained set of truth tables.

54. **(Amended)** A system for quantifying the relatedness of a set of genes, the system comprising:

a multivariate nonlinear predictor constructor operable to construct a multivariate nonlinear predictor based on gene expression level observations for a plurality of candidate genes; ~~and~~

a multivariate nonlinear predictor tester operable to test the effectiveness of the multivariate nonlinear predictor to ~~quantify~~ **generate a quantification indicating** relatedness for the plurality of candidate genes**; and**

a results presenter to present the quantification as indicating relatedness for the plurality of candidate genes.

55. **(Canceled)**

56. **(Canceled)**

57. **(Canceled)**

58. **(Canceled)**

59. **(Canceled)**

60. **(Canceled)**

61. A computer-implemented method of ranking the relatedness of a plurality of genes based on gene expression level observations associated with the plurality of genes, the method comprising:

based on the gene expression level observations, constructing a plurality of multivariate nonlinear predictors to predict the expression of a plurality of target genes out of the genes, wherein the multivariate nonlinear predictors comprise predictive elements comprising an observed gene, thereby associating the multivariate nonlinear predictor with the target gene and at least one observed gene;

testing effectiveness of the plurality of multivariate nonlinear predictors in predicting gene expression to quantify gene relatedness between the genes associated with the predictors by estimating a coefficient of determination; and

displaying a ranked list of gene relatedness among the genes as determined by testing the plurality of multivariate nonlinear predictors.

The following new claims have been added:

62. **(New)** The method of claim 1 wherein the gene relatedness indicates relatedness within a network controlling gene expression.

63. **(New)** The method of claim 1 wherein the gene relatedness indicates relatedness based on chains of interaction among various mechanisms.

64. **(New)** A computer-implemented method for analyzing a plurality of candidate genes for which a plurality of gene expression level observations have been collected to determine which out of the genes are related, the method comprising:

for a plurality of selected permutations of the plurality of candidate genes, performing (a)-(c):

(a) based on data comprising the plurality of gene expression level observations for the plurality of candidate genes, constructing a nonlinear model predicting gene expression for the permutation of the plurality of candidate genes;

(b) predicting gene expression with the nonlinear model; and

(c) measuring effectiveness of the nonlinear model in predicting gene expression, the effectiveness being a quantification of gene relatedness for the plurality of candidate genes of the permutation; and

presenting at least one of the permutations of genes as related and a rating indicating gene relatedness for the permutation.

65. **(New)** The method of claim 64 wherein the gene relatedness indicates relatedness within a network controlling gene expression.